

ABSTRACT

1 A system and method automatically generate an on-line document from raw text into an
2 engaging, interactive form for a plurality of viewers. Unstructured articles are read from an
3 information feed. A computation process extracts and tags proper names of people, products,
4 organizations, and places and categorizes them. An image database is used to link these proper
5 names with image files. The image database consists of a series of attribute-value pairs for active
a 6 searching of names. A URL query string is ^{input}inputted to the database to extract the location of the
7 image in the database file system. An Extensible Markup Language (XML) file is created from
8 the raw text of the article, the list of proper names in the processed data and the image file
9 references. The XML file is stored in a file system. An Extensible Stylesheet Language (XSL)
10 file provides templates containing computational relationships between the text and images. The
11 XML and XSL style sheets are combined to generate a Hypertext Markup Language (HTML)
12 file containing an on-line story of the unstructured articles in a Java Applet which allows the
13 system to provide a variety of interactive behaviors for a final presentation available by a viewer
14 from a browser.